Empowering Organizations to Analyze their Massive Data Stores

Ami Gal
Co-Founder & CEO

SQream

Contact:
CEOCFO Magazine
570-851-1745

Interview conducted by:
Lynn Fosse, Senior Editor
CEOCFO Magazine

CEOCFO: *Mr. Gal, what was the vision behind SQream and where are you today?*
Mr. Gal: When we started, we saw evidence that companies are moving to a data- and insight-oriented world. We thought there would be a need to change the way we analyze data. Data sets were exploding, existing technologies were not able to scale to the magnitude that we thought they would, and organizations were living or dying based on being able to get insights from the data. Algorithms had to consume a lot of data in order make a real difference.

This is where we came up with the idea of using GPUs (Graphics Processing Units) to crunch massive amounts of data in SQL. There were a few things that we did not anticipate, because what happened was data sets were growing even more rapidly than we had expected in the wildest scenarios. Companies were hitting the wall when they tried to make sense of these massive amounts of data – and they absolutely needed to make sense of it. We came prepared to a market that was ready for this kind of solution, even before we had anticipated it.

CEOCFO: *Why is it so hard to process a lot of data?*
Mr. Gal: In order to process so much data you need a lot of storage – fast storage, or a fast way to get data into these kinds of storage. You need to have a lot of compute, a very fast network between all of the nodes, and you need more and more machines in order to process that data. When you get more machines, the network becomes bottomless, because you need to move data between the nodes if you want to do complex squaring. You find yourself brute forcing compute and storage to solve very large and complex problems. Then you find yourself with a thousand nodes, and you find yourself moving a lot of data between these thousand nodes. It works well, but then someone changes the query, or there is another insight, or there is another dimension to that insight they want to add, and all of the optimization you did is not good enough. This happens on a daily basis, when someone wants to get some new insights or change the way they get the insights.

We come at this challenge with the idea that you can use GPUs to create massive processing power in a single node. We can put four GPUs on one node and have twenty thousand computing cores, and if they are efficient at crunching data, you have an entire data center on one node. This is why it's very difficult and why we bring something new that can solve that problem.

CEOCFO: *What is involved in an implementation?*
Mr. Gal: Usually, once we understand the challenge that an organization is facing and how we can help with it, we focus on that and we then can understand the business impact. If there is no business impact, then we do not do an

implementation, because we only like to work with real problems. When we have a business owner that really needs a solution, we either continue it on the cloud or on-prem, and start ingesting the data that is relevant to the problem.

Our ingestion rate is extremely high, even when a lot of terabytes are available in a short time frame. It can be a few minutes or it can be a few hours, and we have the data that you need available for us to ingest. Then we start to run queries that challenge the customer. I would say that customers see value sometimes after a few hours, sometimes after a couple of days, and they already begin to be able to do stuff they could not do before and see insights that they had never seen before. This is a "wow" moment for our customers. This is the place that they understand that they can do everything they dreamt of but could not do before because of the size and complexity of the data and the query. Usually after something between a few hours to a couple of days, they see the light and they like it!

CEOCFO: *Would you give us an example from a couple of different industries of what a company might be looking for and what they can do once you have solved the problem for them?*
Mr. Gal: I'll mention a couple of use cases that may be of interest. The first one comes from the financial world and specifically from fraud detection and credit card or payments. One of the most important things for a payment company or a credit company is to understand a transaction as soon as possible, is it fraudulent or not? This is a well-known problem. Plenty of technologies and every single credit card or payment company is occupied with this issue all the time, every day, throughout the life cycle of the company. Therefore, they invest a lot of time in buying technologies and in creating algorithms, data learning, machine learning, and more algorithms based on the data sets that they have. They have a lot of historical transaction data. They prefer algorithms that run in real-time. Once the transactions are made, they need to know in a fraction of a second if it is accurate or not.

> **"We come at this challenge with the idea that you can use GPUs to create massive processing power in a single node. We can put four GPUs on one node and have twenty thousand computing cores, and if they are efficient at crunching data, you have an entire data center on one node. This is why it is very difficult and why we bring something new that can solve that problem." Ami Gal**

We come in and help these companies as they develop, test, mature and optimize the algorithms. One of the biggest problems of companies that are trying to mature any AI algorithm is the size of the data set on which they run the algorithm and train the algorithm to learn from to optimize the fraudulent situation, as in this case. There are many other cases. We came to one of these well-known worldwide global credit card companies and they were hitting the wall on a daily basis with the size of data sets that they could train, test and optimize the algorithm on. They were bringing in more and more data sources and more and more insights in order to optimize that algorithm, and all of the algorithms needed to be tested and trained and run by zillions of records. They had problems because they needed to update the algorithm every few minutes, every day, every few hours. However, in order to optimize it to the best accuracy, they had to run for a month or it would not work. So we helped them reduce this time frame to between twenty minutes and one hour on a very large broad-based data set. This raised the accuracy of identifying fraud from 89-percent to 97-percent. This was huge for them.

CEOCFO: *Would you give us one more use case?*
Mr. Gal: We have a new reality that includes COVID-19. Again, this is more speaking to a government planning healthcare situation. And now that they are considering re-opening activity or releasing from quarantine, cities would like to see the risk they are taking and how COVID-19 may spread once they open public transportation. This is a very, very large data set of people that includes locations and infection status, what the impact is of opening public transportation, where the specific lines go, and so on. They want to predict in one day, two days, three days, one week, what would be the risk, what is the spread, and what is the best way to open public transportation and still keep the public safe. This involves huge data sets and very complex queries that takes ages. The cities need it right now.

This is another example of how we have been helpful, helping public safety organizations, healthcare situations and city planning bodies that help to fight the COVID-19 proactively. We need to balance the economy by opening the things for the public, but on the other hand we need to keep them safe. Figuring out what is the less risky situation is very important right now.

CEOCFO: *What is the competitive landscape for SQream? What sets SQream apart from others?*
Mr. Gal: There is the addressable market that we go after and this is one thing that we see, there is a technology landscape that we are part of and there is the platform the customer wants to get these insights there are not necessarily the same thing. Usually, when we go to a customer, we offload something from the data source, and this can be Teradata, Exadata or other data warehouses. It is not really that we compete with them, it is more that there are things they cannot do. They just cannot scale to this size or magnitude of data, while we can. So we bridge the gap. In many cases there is a very good partnership between us and these players.

Another situation where we like to be involved is when companies create very large data stores on top of Hadoop This is too much data to have a decent performance or to have performance at all. So they work on analytics with us, on top of any type of solution. We see that a lot on-prem. On the cloud, we definitely see competition. We've just started on the cloud, so we are a new contender and what we see there is usually Amazon AWS, which appears to be a very popular solution. We are definitely going to compete with these guys on their large, high-end customers.

On the technology side, we do not see them too much in the marketplace, because we are focused on different use cases. However, if you look at GPU databases, you can see mostly Kinetica and OmniSci as leaders in this technology category. We see them being pretty active in the e-memory space and the smaller data set. That is the biggest differentiator between us and all the other players, so we definitely see them in dashboards for smaller data sets and with visualization for smaller data sets. We only focus on large data sets, so we fit beyond the 10 to 20 terabytes. It is good for to us and the more the merrier. If you have 100 terabytes or 300 terabytes it is even better. We do not see many of these players working with very large data sets.

CEOCFO: *Would you tell us about the recent funding and how that will move SQream ahead?*
Mr. Gal: We have been fortunate to grow quite a bit in the last couple of years in terms of business. Our growth rate is pretty high and the demand for a solution is also growing. We need to scale as a company and as a team in order to support the types of customers and the types of big data challenges that we are facing. Customers and data sets are now growing to beyond 10-thousand terabytes and we have customers that are starting to use us for data sets that are bigger than 10 terabytes. One side of the round was to double down on growth with a go-to-market team and to support our global customers.

The other side that is just as important, is the cloud. We think that our customers and the market customers want to analyze raw data sets wherever they are. It can be on-perm, it can be on the cloud and it can be on the edge. The main focus we have right now is in launching our product from the cloud and providing our product as a service. We think we will need to invest a lot in the go-to market area and make sure our customers are happy with our solutions on the cloud and on the edge. This is more about going after the growing segment of the same type of customer we are helping on-prem, who are moving to the cloud. These are the main reasons for the round and because we are growing quite a bit. We have also been fortunate to have quite a bit of demand in this round.